

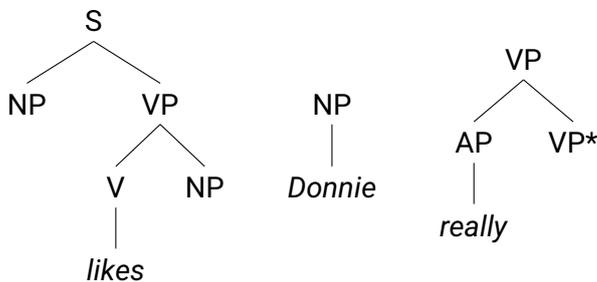
Grammatikimplementierung mit TAG

Laura Kallmeyer, Benjamin Burkhardt

Wintersemester 2017

Übung 1 (Besprechung 19.10.2017)

(1) Sei G_{LTA} eine LTAG bestehend aus folgenden Bäumen:



- a. Welche Elemente sind Mütter- und Tochterknoten, terminale und nicht-terminale Symbole, Fuß- und Wurzelknoten und welche sind Blätter? Welche Bäume sind Elementar-, Initial-, Auxiliar-, Substitutions- und Adjunktionsbäume?

Lösung:

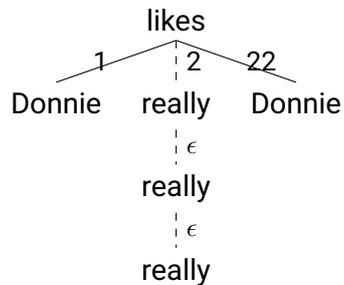
- Jeder Knoten, der andere Knoten direkt dominiert, ist ein Mutterknoten.
- Alle Knoten, die von einem anderen Knoten direkt dominiert werden, sind Tochterknoten des dominierenden Knotens.
- Die lexikalischen Elemente sind terminale Symbole. Alle anderen Symbole, mit denen die syntaktische Kategorie eines Knotens bezeichnet wird, sind nicht-terminale Symbole.
- Der Wurzelknoten ist der höchste Knoten eines Baums.
- Der mit einem * markierte Knoten eines Adjunktionsbaums ist der Fußknoten.
- Alle Knoten, die keine weiteren Knoten dominieren, sind Blätter.
- Alle nicht abgeleiteten Bäume sind Elementarbäume.
- Bäume mit dem Startsymbol S im Wurzelknoten sind Initialbäume.
- Substitutions- und Adjunktionsbäume sind Auxiliarbäume.
- Substitutionsbäume sind die Bäume, die die für Substitution markierten Blätter eines Initialbaums ersetzen können.
- Fußknoten und Wurzelknoten eines Adjunktionsbaums haben die selbe syntaktische Kategorie. Der Fußknoten ist mit einem * markiert. Adjunktionsbäume erweitern die Struktur eines Initialbaums.

- b. Welche Sprache wird G_{LTA} generiert? (Notieren sie n Wiederholungen eines beliebigen Strings a als a^n .)

Lösung: $S_{G_{LTA}} = \{Donnie\ really^n\ likes\ Donnie\}$

- c. Bestimmen sie den Ableitungsbaum für den Satz mit sechs terminalen Symbolen!

Lösung:



- (2) Sei G_{CF} eine kontextfreie Grammatik mit den folgenden Ersetzungsregeln:

S	→ NP VP	NP	→ Paul
VP	→ V NP	NP	→ football
V	→ plays		

- a. Der Satz *Paul often plays football in the garden* kann nicht durch G_{CF} generiert werden. Erweitern sie die CFG sinnvoll, um den Satz generieren zu können. Jedem Einzelwort sollte eine lexikalische Kategorie zugeordnet werden können.

Lösung:

S	→ NP VP	VP	→ V NP	AdvP	→ Adv
NP	→ Paul	V	→ plays	Adv	→ often
NP	→ football	VP	→ AdvP VP	PP	→ P NP
NP	→ Det N	VP	→ VP PP	P	→ in
Det	→ the				
N	→ garden				

- b. Geben sie mindestens einen Satz an, der durch G_{CF} lizenziert wird, der jedoch nicht von Muttersprachlern geäußert werden würde. Welche Vereinfachung in G_{CF} verursacht diese Übergenerierung?

Lösung:

Football plays football.